Exhibit 6

The Causal Relationship Between Prenatal Acetaminophen Use, Neurodevelopmental Disorders (NDD), Attention-Deficit/Hyperactivity Disorder (ADHD), and Autism Spectrum Disorder (ASD).

Rule 26 Rebuttal Report of Andrea Baccarelli, MD, PhD, MPH

I have reviewed the July 21, 2023, expert reports from Defendants' experts and provide the following rebuttal report. My rebuttal report focuses on the most fundamentally incorrect assertions and opinions from Defendants' experts but is not intended to be an exhaustive summary of the inaccuracies or incorrect opinions set forth in Defendants' experts' reports. My affirmative report contains my responses to many such inaccuracies and incorrect opinions, and it is incorporated here by reference. I primarily focus my rebuttal on the opinion of Dr. Pinto-Martin, but my rebuttal points are responsive to most of Defendants' expert reports, as many echo the same, incorrect points. The facts and data I have considered in forming my opinions in this case are referenced in my report, this rebuttal, and on the List of Materials Considered. I also rely on and incorporate the rebuttal reports of Dr. Robert Cabrera, PhD (teratology), Dr. Brandon Pearson, MSc, PhD (toxicology), Dr. Eric Hollander, MD (neurodevelopment) and Dr. Stan Louie, PharmD (pharmacology) to rebut Defendants' expert reports that target my opinion.

I. Negative control exposure in acetaminophen studies of ADHD, ASD, and other Neurodevelopmental Disorders

Dr. Pinto disagrees that many relevant studies have ruled out genetics and other (highly unlikely) residual confounding through the use of negative control exposures (NCEs) and argues that my explanation of the findings from studies using NCEs "is flawed." (Pinto-Martin Report at 98). This is not correct, as explained below.

a. Dr. Pinto-Martin ignores statements from the study authors themselves regarding negative controls

To begin with, her argument is inconsistent with several statements of the study authors themselves, who viewed the negative-control exposures as compelling evidence that the association between prenatal APAP exposure and neurodevelopmental disorders is *not* the result of some unidentified form of confounding. For example, the Ystrom authors note that "maternal preconceptional use"—*i.e.*, a woman's use of Tylenol before she became pregnant—was highly correlated with a woman's use during pregnancy, thus supporting "maternal preconceptional use as a negative . . . control."[1] Indeed, the authors went on to say that this result was "consistent with a causal link," *i.e.*, *not* consistent with some unidentified confounding. The Liew 2019 authors stated that the "the absences of associations with acetaminophen use in the pre-pregnancy and post-pregnancy periods are the true NCE tests, and they suggest that variables that do not vary over the years—such as genetics, maternal chronic diseases, or socioeconomic status—do not explain the association observed for acetaminophen exposure at the time of pregnancy."[2] The Liew 2019 authors went on to conclude that "[t]he findings of our NCE analyses corroborate those of prior reports suggesting that prenatal acetaminophen exposure may influence neurodevelopment," *i.e.*, suggesting a causal link. The Stergiakouli authors stated that their negative control analysis—*i.e.*, the fact that the "associations were not observed for postnatal or partner's acetaminophen use"—provided substantial evidence against confounding and provided support for "an intrauterine mechanism," *i.e.*, causation.[3] Dr. Pinto-Martin provides no compelling explanation as to why the study authors *themselves* are mistaken on the usefulness of these negative controls, or why she is in a better position than the authors to evaluate the strength of the negative controls.

b. Dr. Pinto-Martin ignores the weight of the data on negative controls

In any event, on this issue, the data speaks for itself. Below, I provide a table with all the studies that report NCEs related to prenatal acetaminophen use and neurodevelopmental disorders (NDDs).

Table. Studies of the association between acetaminophen use during pregnancy and neurodevelopmental conditions that include negative control exposure analyses.

Outcome	Study	Effect of maternal acetaminophen use during pregnancy	Negative Control Exposure	
			Type	**Effect estimate**
ADHD	Ystrom, 2017	HR: 1.20 (95% CI: 1.02, 1.24) (ever use)	Maternal acetaminophen use six months before pregnancy	HR: 0.95 (95% CI 0.85–1.06)
			Paternal acetaminophen use 6 months before pregnancy	HR: 1.27 (95%CI 1.08, 1.49)
ADHD	Stergiakouli, 2016	RR=1.43 (95% CI 1.18-1.73)	Maternal postnatal acetaminophen use 61 months postnatally	RR: 1.10 (95% CI 0.89-1.36)
			Partner's acetaminophen use 61 months postnatally	RR: 1.27, 95% CI: 0.96-1.69
ADHD	Liew, 2019	OR = 1.34 (95% CI: 1.05, 1.72)	Maternal acetaminophen use 4 years before pregnancy	OR: 1.12 (95% CI: 0.91, 1.38)
			Maternal acetaminophen use 4 years after pregnancy	OR: 1.05 (95% CI: 0.88, 1.26)
ADHD	Chen, 2019	OR = 1.20 (95% CI 1.01, 1.42)	Maternal acetaminophen use three months before pregnancy	OR: 1.06 (95% CI 0.90, 1.25)
Other neuro-developmental outcomes	Tronnes, 2020	Communication problems: RR: 1.18 (95%CI 0.86, 1.60) Externalizing problems: RR: 1.22 (95%CI 0.93, 1.60) Internalizing problems: RR: 1.36 (95%CI 1.02, 1.80) Emotionality: Beta: 0.13 (95%CI -1.08, 1.33) Activity levels: Beta: 0.51 (95%CI -0.57, 1.60) Sociability: Beta: -0.07 (95%CI -1.02, 0.88) Shyness: Beta: -0.24 (95%CI -1.27, 0.80)	Maternal acetaminophen use six months before pregnancy.	Communication problems: RR: 1.19 (95%CI 1.02, 1.38) Externalizing problems: RR: 0.99, (95% CI 0.87, 1.14). Internalizing problems: RR: 0.96 (95% CI 0.84, 1.10). Emotionality Beta: 0.36 (95% CI -0.08, 0.81) Activity levels Beta: -0.80 (95% CI: -1.23, -0.36) Sociability RR: 0.22 (95% CI -0.22, 0.66). Shyness: RR: 0.35 (95% CI -0.10, 0.80).

HR: Hazard ratio

This table demonstrates a clear, statistically significant association between maternal acetaminophen and NDDs in most studies for the effect of maternal acetaminophen use during pregnancy on neurodevelopmental outcomes (see third column from the right—and especially results marked in red). The negative control analyses are entirely consistent with those associations being causal because the negative control (the variable we *would* expect to be correlated with NDDs if the association was due to some sort of confounding) was almost uniformly *not* statistically significantly associated with NDDs. This is an extremely persuasive set of results suggesting that the link between prenatal APAP exposure and NDDs is indeed a causal one. Overall, the studies above present 14 separate analyses that use negative controls. Eleven of those results fully support the absence of any residual confounding and are reported in red in the table. As discussed below, the remaining three results—reported in green—do not provide a compelling reason to disregard the obvious weight of the overall negative control results.

Across studies, any effect of residual confounding, if present, would be expected to be similar to that found for the NCEs, *i.e.*, it would produce findings with adverse, protective, or no effects depending on the population investigated. For example, if there really *was* some set of genes that was correlated both with a woman's propensity to take APAP generally and her propensity to have a child with a NDD, we would expect that a woman's APAP use before and after pregnancy would be correlated with her likelihood to have a child with an NDD. But as the table above makes clear, that is not the case in the majority of the studies. This stands in stark contrast to the overall evidence on the link between prenatal APAP exposure and NDDs, as the same studies consistently show significant associations between prenatal acetaminophen exposure and ADHD, ASD, or other NDDs.

c. Even the scant results in which negative controls showed an effect do not provide persuasive evidence of confounding

In Dr. Pinto-Martin's report, she does not present the full analysis of all negative controls that I present above. Instead, she focuses only on a few results that (in her view) suggest residual confounding. But even among her cherry-picked set of results, not all of them are consistent with the presence of confounding that would explain the results.

For one example, in Trønnes 2020[4], prenatal maternal acetaminophen during pregnancy has a *positive* association with higher activity level, while the negative control (*i.e.*, maternal acetaminophen six month before pregnancy) has a *negative* association with activity level. In other words, the primary analysis and negative controls show opposite results. At most, this result suggests that if confounding existed, it would be toward the null, thus leading to an underestimation of the true effect of the exposure during pregnancy. This does not suggest the kind of residual confounding that could explain the association between prenatal APAP exposure and NDDs.

For another example, in Ystrom 2017[1], the authors used paternal use of acetaminophen 6 months before the pregnancy as the negative control. This negative control is as tangential as they come, looking at whether a *father's* APAP use was associated with a child's likelihood of developing ADHD. Indeed, as recognized by the Ystrom authors, a father's APAP use is not an ideal negative control, because we now know that many exposures, including acetaminophen, can modify DNA methylation systemically and in the male testis, an epigenetic mark that can be transmitted to the offspring and mediate the causation of neurodevelopmental disease.[5,6] In other words, the fact that a father's use of APAP is associated with NDDs does not clearly suggest residual confounding. Instead, that association may simply suggest an additional pathway by which acetaminophen may have an adverse effect on a child's neurodevelopment.

d. There is no evidence that pre- and post-pregnancy use are invalid negative controls

Although Dr. Pinto-Martin states that pre- and post-pregnancy use is not a valid control because women's APAP habits might change during pregnancy, that makes her novel theory of genetic confounding even more implausible: It assumes that there exists a gene that (a) does make a woman more likely to take APAP while pregnant; (b) does not make a woman more likely to take APAP before pregnancy; (c) does not make a woman more likely to take APAP after pregnancy; and (d) causes NDDs in the fetus. Suffice it to say that there is no evidence suggesting that this kind of genetic confounder exists.

Moreover, Dr. Pinto-Martin seems to believe that paternal/partner acetaminophen use is the only valid negative control, but her arguments in support of that belief are contradictory. Dr. Pinto-Martin seems to indicate instead that the paternal/partner's use of acetaminophen, as reported in Ystrom 2017 and Stergiakouli 2016, would be a valid NCE but that the maternal use of acetaminophen during pregnancy is likely to be different from maternal acetaminophen use (by the same women) before or after pregnancy and thus not a valid NCE. It is counterintuitive that Dr. Pinto-Martin is concerned about different use in the same women over time, while believing at the same time that the use by a partner/father would follow similar patterns as in their female partners. To the extent that there are material differences between a pregnant woman's propensity to take APAP and *that same* woman's propensity to take APAP while not pregnant, there is every reason to think that there would be even greater differences between a pregnant woman's propensity to take APAP and her partner's propensity to do the same. It is difficult to explain Dr. Pinto-Martin's rationale to credit paternal APAP use (but not pre- and post- maternal APAP use) as a valid negative control and suggests results-oriented reasoning on her part.

e. The magnitude of results in the negative controls is not enough to explain the association between prenatal APAP exposure and NDDs

In any event, assuming that Dr. Pinto-Martin is correct about which controls are valid, the math still does not add up. Even if pre-natal *paternal* APAP use were a valid NCE variable, that would mean that the results from that NCE variable would represent the *total* overall strength of *all* unmeasured confounding, including but not limited to genetics, that might influence the association between maternal pregnancy acetaminophen use and ADHD. And to put it bluntly, that result is not nearly large enough to explain the effect overserved here. Specifically, the observed odds ratio for paternal APAP use is much smaller than the odds ratio (of 3.0) that Ricci et al. have estimated to be needed for unmeasured confounding to explain the association between maternal pregnancy acetaminophen use and ADHD.[7] In short, even if paternal use were a compelling negative control, it in no way suggests that *all* of this observed association is due to some as-yet-unidentified confounder—from genetics or any other source.

f. Conclusion on negative controls

In conclusion, the evidence from the NCEs and the consistent findings from the main analyses of various studies support the strong association between prenatal acetaminophen use and ADHD, ASD, or other NDDs. The table of NCEs clearly shows that residual confounding, if any, cannot explain away the significant associations observed in the main analyses. The authors who used NCEs employed a variety of different variables, hence adding to the robustness of my conclusions. This collective evidence strengthens the validity and credibility of the reported associations between prenatal acetaminophen exposure and the neurodevelopmental disorders of ASD and ADHD, debunking the claims of the comment provided by Dr. Pinto-Martin.

II. Sibling-controlled studies

a. The findings of the sibling-controlled design are severely biased toward the null and misinterpreted

Dr. Pinto-Martin disagrees with the mathematically demonstrated tenet that sibling-controlled studies adjust for both confounding and intermediate variables, an opinion also brought forward by Dr. Faraone and Dr. Chung. However, the same authors that published the sibling-controlled studies of acetaminophen[8]—*i.e.*, the same authors on which Drs. Pinto-Martin and Faraone rely— state explicitly this exact same point in their paper—that sibling-controlled studies do take away the effect mediated by intermediate variables. Nonetheless, Dr. Pinto-Martin argues that only Sjolander[9] has stated that in the literature. She fails to mention, however, that Sjolander provides a *mathematical demonstration* of this concept, a demonstration that I found correct. Dr. Pinto-Martin failed to identify errors in Dr. Sjolander's mathematics or provide her own mathematical demonstration that would justify a contrary conclusion. If an author wrote a paper that demonstrates that 2 + 2 = 4, no one could disprove that paper by saying "it is the only paper demonstrating the validity of that sum." One would have to disprove that 2 + 2 = 4. So it is for Sjolander's mathematics, which Dr. Pinto-Martin makes no effort to dispute, much less disprove.

Dr. Faraone also believes that my interpretation that sibling-controlled studies adjust for effect modifiers is unfounded. The role of intermediate adjustments brought about within sibling-controlled studies is fully sufficient to support my point. Furthermore, effect modifications can create differences across subjects in intermediate responses; therefore, the results of effect modifications will be adjusted for when estimating sibling-controlled estimates.

b. Sibling-controlled studies are not commonly used in epidemiology and are known to be susceptible to bias

Dr. Pinto-Martin also argues that sibling-controlled studies are widely used and accepted in epidemiology. Regardless of how often this kind of study design is employed, two points are worth highlighting. First, across the 30+ studies on APAP and ADHD, ASD, and other NDDs that I have evaluated, only two studies (both from the same group) used the sibling-controlled design. One found an effect difference between exposed and unexposed siblings.[10] One did not.[8] Second, the most recent edition (2021) of Rothman's *Modern Epidemiology*[11]—widely considered the definitive reference textbook in epidemiology—dedicates less than a page of commentary to "Sibling Controls." And most of that commentary is about the flaws and limitations of sibling control studies. As Rothman and his co-authors note, many of those flaws and limitations would bias any results toward the null, adding another reason why this study design is likely to underestimate effects. For total clarity on this subject, I report the entire text on Sibling Controls by Rothman below:

> *<u>Sibling Controls</u>*
> *Matching study subjects with siblings may be used in both cohort and case-control studies to control for shared genetic and environmental factors. For example, in a case-control study of breast cancer that included both sister controls and controls sampled from the source population, the two sets of controls differed substantially with regard to many established risk factors for breast cancer, and estimates of associations for established breast cancer risk factors were nearer to expectation using sister controls than population controls. Sibling case-control designs can reduce confounding by shared factors, but they can also increase nondifferential and differential exposure misclassification and potential confounding by nonshared factors. For example, imagine a case-control study of oral clefts in which maternal smoking during pregnancy*

was compared between case infants and their nonmalformed siblings. Within a sibling pair discordant for maternal smoking, the likelihood of the apparently unexposed pregnancy to be actually exposed (e.g., smoking not recorded for one of the pregnancies) would be higher than the overall underrecording probability in the general population of pregnancies because the mother smoked at one point. In addition, the fact that the exposure differs between siblings may be associated with a factor that in turn is associated with the outcome. The discordance for within-sibling confounders would be larger than the discordance when cases are compared with random persons in the population, so the association between the confounder and exposure would be stronger in within-sibling comparisons than in other case-control sampling strategies. If young mothers quit smoking before subsequent pregnancies, smoking would be associated with younger age within-sibling comparisons. If young age were the strongest risk factor for oral clefts, the sibling-matched design could be more biased than a nonsibling approach unless age were controlled in the analysis.

Contrary to the suggestion of Drs. Pinto-Martin and Faraone, the question is not whether sibling control methods are common. Instead, the question is whether they provide—as used in the case at hand—information that proves or disproves the overwhelming evidence that acetaminophen causes ADHD, ASD, and other NDDs. Any pharmacology and toxicology textbook contains hundreds of pages showing that medications and chemicals activate a series of intermediate steps leading to toxicity. Therefore, most of the adverse effects of a medication or a chemical, including acetaminophen might well be *indirect, i.e.*, mediated by those steps. This information, compounded by the remarks by Rothman that sibling controls can additionally introduce bias, indicate that the sibling-controlled studies estimate only a small part of the effect (direct effect) and that even this effect might be biased toward the null. In light of these limitations, it is all the more compelling that the Brandlistuen study showed "in a sibling controlled analysis" that "children exposed to long-term use of paracetamol during pregnancy had substantially adverse developmental outcomes." In other words, this study showed that siblings who *were* exposed to APAP while in utero had substantially worse developmental outcomes than their siblings who were *not* exposed to APAP while in utero. That result is entirely consistent with a causal relationship, and not at all consistent with the theory espoused by Dr. Pinto-Martin and Dr. Faraone that these associations are due to some as-yet-unidentified confounder, whether genetic or otherwise.

c. Gustavson 2021 – sibling-controlled study with a very small sample size

Gustavson et al. (2021)[8] reported that long-term exposure (29 days or more) to prenatal acetaminophen was associated with a two-fold increase in risk of ADHD diagnosis (adjusted HR = 2.02, 95% C.I = 1.17–3.25). Dr. Pinto-Martin points to the sibling controlled-analysis from Gustavson 2021 as evidence that any causal association based on prenatal use of acetaminophen is attenuated and is actually explained by genetics. (Pinto-Martin at 96). She also criticizes me for not assessing the sibling-controlled aspect of Gustavson 2021 for my dose response analysis. She is wrong on both points as explained below.

As for whether the sibling-controlled analysis suggests confounding by genetics, I have already explained at length why sibling control studies likely underestimate the true risk from APAP. (Baccarelli Report at 117-122). I will emphasize one additional feature here: sibling-controlled studies reduce the overall sample size, and thus reduce the power available to detect a signal. The entire sample size for Gustavson 2021 was 26,613, but as I noted in my Appendix for my Navigation Guide analysis, for the sibling-control analysis, "[o]nly discordant siblings contributed to power; siblings were discordant on exposure for 29 days and the outcome in **34** families." In other words only n=34 pairs contributed to the analysis of 29+ days, a sample size that is hugely underpowered. Even the authors of the study note, "As only discordant siblings

contribute to information in sibling control models, even the current very large birth cohort provided limited statistical power. Hence, the results need to be replicated in other studies." (Gustavson 2021 at 8) As Gustavson reports a prevalence of ADHD equal to 2.8% in their cohort, applying that proportion to n=34 families would bring us to the conclusion that the analysis is based on only approximately 2-3 cases of ADHD, a number that clearly does not allow to draw any conclusion. In fact, the Gustavon 2021's own authors agreed with my position and stated "[a]s only discordant siblings contribute to information in sibling control models, even the current very large birth cohort provided limited statistical power. Hence, the results need to be replicated in other studies."[8] The appendix to the study further underscores the limited power of the sibling-control analysis, as the study only had 306 mothers who had children discordant on both exposure and outcomes. Since these 306 pairs are the only informative data for the sibling-controlled part of the study, it is even smaller than Ji et al.[12] and Baker that Dr. Pinto-Martin criticized for their purportedly limited power (Pinto-Martin at 81), in spite of the huge advantage over Gustafson 2021 of having objective measurements of exposure that reduce exposure misclassification, therefore increasing—given a set sample size—their capacity to identify true effects. It is difficult to understand how Dr. Pinto-Martin maintains that the small sibling-controlled analysis in Gustavson 2021 might disprove a comprehensive body of literature that features much larger studies, including a few that each include >50,000 children and thousands of ADHD cases. The reduced power of the sibling-controlled analysis in Gustafson 2021 further decreases the likelihood that the sibling-controlled analysis would detect a significant association.

As for dose response, Dr. Pinto-Martin says that "Gustavson et al. 2021, found no dose-response relationship between maternal acetaminophen use and an increased risk of ADHD." (Pinto Martin Report at 104). She goes on to criticize me for (as she puts it) failing to "mention this study when discussing dose-response." (Pinto Martin Report at 105). Both of these statements are demonstrably false. First consider her assertion that "Gustavson et al. 2021, found no dose-response relationship." (Pinto Martin Report at 104). The authors of Gustavson 2021 would surely disagree, since they state that there *was* a "dose response association[] between number of days of acetaminophen exposure and risk of ADHD" in their top-level analyses—and they demonstrate as much in Table 2 of the paper itself. Dr. Pinto-Martin is also mistaken that I failed to "mention this study when discussing dose-response." I specifically discuss Gustavson 2021 in my Navigation Guide analysis. Indeed, I specifically discuss *both* the top-line results *and* the sibling-controlled results.

What Dr. Pinto-Martin really seems to be suggesting is that I should have relied *exclusively* on the sibling-controlled results from Gustavson 2021—for dose response and everything else—and blinded myself to the top-line results. But there is no reason to exclude the other results from that study from my analysis—I considered all of the evidence, including both the top-line results and the sibling-controlled analysis, as a good epidemiologist should. And to the extent that I placed less weight on the sibling-controlled component of Gustavson 2021, it is for the reasons given above and in my original report.

III. Defendants' experts misinterpret and overstate heritability's role in the development of neurodevelopment disorders

a. The Misinterpretation of Heritability: Understanding its Limitations in Neurodevelopmental Disorders and Environmental Toxicity

Dr. Pinto-Martin, as well as Defendants' other experts, assert that the neurodevelopmental disorders of ASD and ADHD are attributable to heritability or genetics, but their opinions are premised on a number of

incorrect assumptions. I outline those below, but I also refer to and incorporate the rebuttal report of my colleagues, Dr. Robert Cabrera and Dr. Eric Hollander.

As a general matter, Dr. Pinto-Martin's statement that genetics is the predominant factor in ADHD and ASD etiology is incorrect. (Pinto-Martin Report at 27). It stems from a misunderstanding or misrepresentation of the concept of "heritability" and the roles of genes and environment in the etiology of complex diseases, including ADHD and ASD.

Dr. Pinto-Martin states in her report that "[r]esearch has shown that genetics are the predominant cause of ASD." (Pinto-Martin Report at 27). She bases that statement on data from recent "large-scale twin studies [that] have reported that the heritability of ASD—in other words the percentage of ASD cases attributable to inherited genetic factors-rather than either environmental factors or random chance—ranges from 80 to 90%". (Pinto Martin Report at 27 (citing Tick et al. 2016)). Dr. Pinto-Martin makes similar statements regarding ADHD, including "[g]enetics are the predominant factor in the etiology of ADHD, with studies reporting heritability estimates of approximately 75%." (Pinto-Martin Report at 67). Interestingly, Dr. Pinto-Martin cited Tick et al. 2016 for this proposition (Pinto-Martin Report at 27-28) but did not mention the much newer and much larger study by top researchers from around the world published in JAMA Psychiatry, Bai et al. 2019[13]. With respect to the issue Dr. Pinto-Martin addressed, the results from Bai et al. were not dramatically different: "Based on population data from 5 countries, the heritability of ASD was estimated to be approximately 80%." The Bai study, however, examines maternal contribution (*i.e.*, the noninherited genetic influences originating from mothers beyond what is inherited by the offspring) to the development of ASD and finds "no support for the contribution from maternal effect". Bai also concluded that nonshared environmental exposures contributed at most 26.5% to the variability of the risk of ASD and that "nonshared environmental factors also consistently contribute to risk".

I agree with Dr. Pinto-Martin that ADHD and ASD are, in general, heritable conditions. However, as pointed out by a 2017 comprehensive review of ASD heritability, "[e]arlier twin studies suggested heritability as high as 80–90% for ASD with little contribution from the environment. According to recent evidence, up to 40–50% of variance in ASD liability is determined by environmental factors."[14] Therefore, more contemporary estimates put ASD heritability at a much lower level. Also, estimates of ADHD heritability of ADHD show huge variations across studies ranging from as little as 30%[15]--particularly for adult ADHD--to as high as 90%[16]. Therefore, there is little certainty about the degree of heritability of these conditions. However, whatever the heritability of ASD and ADHD might be, the fundamental flaw in Dr. Pinto-Martin's view is her use of the heritability of these conditions to attempt to exclude environmental influences. That argument is completely misguided, as explained below and further explained in Dr. Cabrera's rebuttal report.

Unfortunately, Dr. Pinto-Martin is not alone in making this incorrect assumption. Heritability is a statistical term often misinterpreted to mean that a significant percentage of a trait or disorder is directly and exclusively inherited through genes. Although this concept may have been accepted 50 years ago, modern genetics appreciates that neurodevelopmental disorders such as ADHD and ASD are complex, multifactorial, and resulting from the interplay of both genes and environmental factors. A common metaphor used to describe these type of conditions is that the genes load the gun, but the environment pulls the trigger. This section aims to clarify the misconceptions surrounding heritability and highlight its

limitations in understanding the risk of neurodevelopmental disorders such as ADHD and ASD, as well as its inability to exclude the adverse effects of environmental toxicity on these disorders.

b. High heritability does not exclude the influence of environmental factors

1. **Heritability does not mean inheritability:** As Moore and Shenk[17] point out, there is a common misunderstanding between the terms 'heritable' and 'heritability,' and 'inheritable'. Heritability refers to the amount of variation in a trait in a population that can be ascribed to genetic differences, but it does *not* mean that the trait is directly inherited through genes. Therefore, a high heritability index does not imply that environmental factors cannot play a role in the emergence of a trait or disease. In fact, several theorists suggest that the term "heritability", often confused with inheritability in everyday language, might not be apt for use in human genetics, and propose its discontinuation.

2. **The environment has strong influences on phenotypes with high heritability**: Traits with high heritability are modified by environmental factors. A classic example of this is height, which has a heritability often measured at 90%, a level similar to the highest estimates of ASD and ADHD heritability. However, environmental factors like nutrition significantly affect this heritable trait. For instance, South Koreans are on average nearly five inches taller than North Koreans, despite similar gene pools, likely due to differences in nutrition availability. As this example shows, environmental factors can still influence an outcome even if (and indeed especially if) an outcome is highly heritable.

3. **Misinterpretation of high heritability**: Some traits, like voting behavior (72% heritable) and criminal behavior (45% heritable), have surprisingly high heritability indices. However, these estimates do not exclude the influence of environmental factors. To state the obvious, regardless of one's genetics, someone must still decide who to vote for and whether to commit a crime—and environmental factors influence both of those decisions. Instead, high heritability indices indicate a complex interplay between genes and the environment, where certain genetic predispositions may interact with environmental circumstances to bring about a particular behavior or condition.

Another classic example is that of earring-wearing, which—at least in the early 20th century—used to be 100% heritable since earrings were worn only by people who were female, a sex-specific trait linked to having two X chromosomes. However, changes in societal norms have led to more men wearing earrings, indicating that the environment (in this case, societal trends) can have a significant impact on a seemingly 100% 'heritable' trait.

5. **Heritability estimates can be biased and often inflated**: Heritability estimates are influenced by various factors including population stratification, indirect genetic effects, assortative mating, and other environmental factors. These factors can significantly alter heritability estimates.

6. **Heritability of IQ**: The discussion around the heritability of IQ serves as another example. Some studies estimate IQ's heritability as high as 80%, and on average, IQ scores can vary as much as 15 IQ points between various races. However, it would be a tragic mistake to misinterpret these differences as being driven by genetics, as some have offensively done. Despite the very high heritability of IQ, these disparities between races are more accurately attributed to socio-economic, educational, and other environmental factors—which continue to play a role despite the high heritability of IQ.

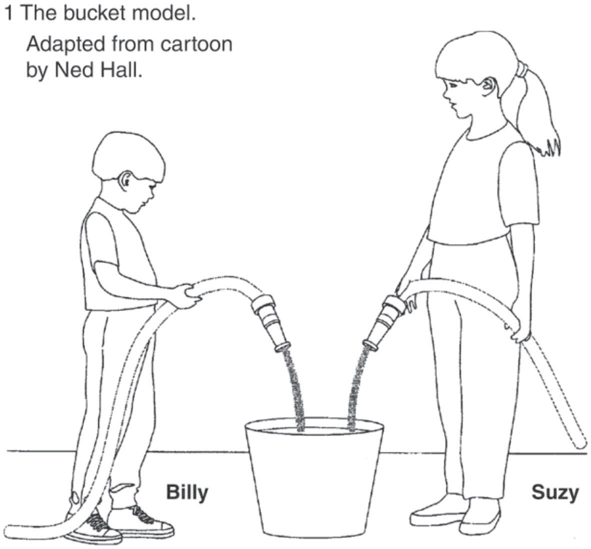
c. Gene-environment interactions are the root cause of neurodevelopmental disease

An assumption commonly made in heritability estimates is that genetic and environmental factors influence traits are independent of each other and mutually exclusive, while in reality, they often interact. For instance, a certain genetic predisposition may only manifest under specific environmental conditions. Ignoring these interactions can lead to an overestimation of the genetic contribution. This is the most important point to consider to accurately interpret heritability estimates and the role of genes and environmental determinants.

In this context, it is clear why it is difficult to accurately capture the biological inheritability of complex traits. The process of trait development is intricate, involving multiple variables and dynamic interactions, particularly in natural settings where environmental factors are less controlled. Consequently, precise predictions about inheritability often remain unachievable except for within experimental or artificial contexts where environmental variables can be strictly controlled. This challenge and the shortcomings of measures of heritability are reviewed in detail by Moore and Shenk. Indeed, heritability statistics retain their relevance in certain contexts, notably selective breeding programs where developmental environments can be closely regulated. However, in uncontrolled environments—including the assessment of the impact of genetics vs environment in ADHD and ASD etiology, these statistics provide limited information.

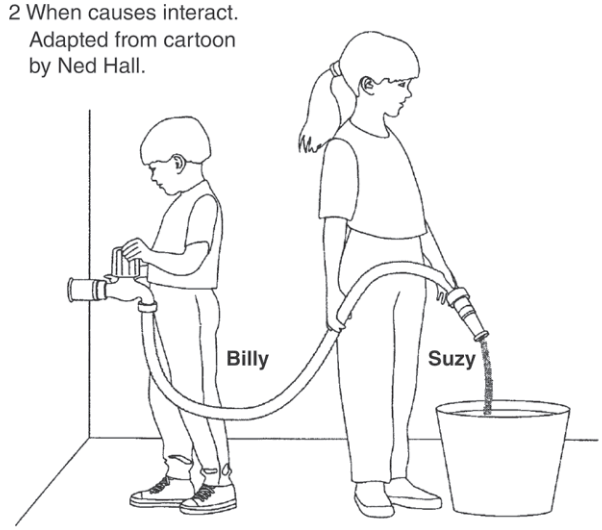
At the same time, all of our characteristics are also influenced by nongenetic factors that interact with the relevant genetic factors. This renders heritability statistics—which are based on the assumption that they are mutually exclusive—not interpretable in terms of gene vs. environment contribution. The Figure below explains why the genetic component and the environmental component of risk are not mutually exclusive.

1 The bucket model.
Adapted from cartoon
by Ned Hall.



Here is a bucket: Billy fills it with 40L of water; then Suzy fills it with 60L of water. So, 40% of the water in the bucket is due to Billy, 60% to Suzy.

2 When causes interact.
Adapted from cartoon
by Ned Hall.



But suppose instead that what happened was this: Suzy brought a hose to the bucket, and then Billy turned the tap on. Now how much of the water is due to Billy and how much to Suzy?

Answer: The question no longer makes any sense.

d. Lack of genes responsible for ASD and ADHD.

Dr. Pinto-Martin's report fails to take into account the evidence discussed above and rather presents an over-simplified, inaccurate understanding of the isolated influence of genes on ASD and ADHD. Her inaccurate portrayal of genetics is confirmed by the lack of consistently identified genes responsible for these disorders. For instance, a recent umbrella review by Qiu et al., 2022[18] performed a comprehensive review of all studies conducted to identify genes associated with the risk of ASD. To provide a robust synthesis of published evidence of genetic studies for ASD, the authors performed an umbrella review (UR) of meta-analyses of genetic studies for ASD. They systematically searched eight publication databases from inception to March 31, 2022. The authors reviewed the original evidence and classified the strength of the evidence of risk genes for ASD into five levels: convincing (class I), highly suggestive (class II), suggestive (class III), weak (class IV), and not significant (Class V).

Overall, 12 significant SNPs of CNTNAP2, MTHFR, OXTR, SLC25A12, and VDR were identified from 41 SNPs of nine candidate genes in 28 meta-analyses. Of those, only two SNPs were classified in as "associations with suggestive evidence" (class III), while the remaining were categorized as having "associations with weak evidence" (class IV). Notably, no genes were classified as having convincing (class I) or highly suggestive (class II) evidence. Of course, this situation is even more prominent for ADHD, a disorder often considered to have lower genetic background than ASD.

In conclusion, there is little definitive evidence about genes consistently associated with ASD and ADHD. The particular study highlights the complex, multi-faceted nature of ASD, suggesting that a single or small number of genes may not fully explain the development of this disorder.

IV. Personality and Neuroticism Do Not Explain Association Between Prenatal Use of Acetaminophen and ASD/ADHD

Dr. Pinto-Martin's opinion that the association between prenatal use of acetaminophen and the neurodevelopmental disorders of ASD and ADHD is attributable to confounding by indication is not scientifically supported. (Pinto-Martin Report at 17-21). Dr. Pinto Martin fails to cite a scientific source to support her position that women with mental health issues are more likely to take acetaminophen while pregnant. (Pinto-Martin at 17-21). Lupatteli et al., 2023[19] undercuts Dr. Pinto-Martin's conclusion. That study assessed the association between five maternal personality traits, including neuroticism, and, among other things, prenatal use of acetaminophen. The study found prenatal use of any acetaminophen associated with neuroticism to have a crude odds ratio of 0.99 (0.84-1.17) and an adjusted odds ratio of 1.11 (0.93-1.32) with a 95% confidence interval and extended prenatal use of acetaminophen (*i.e.,* all three trimesters) to have a crude odds ratio of 1.10 (0.92-1.33) and adjusted odds ratio of 1.31 (1.08-1.59) with a 95% confidence interval. The authors concluded "[w]e found that high neurotic disposition in the woman was associated with a modest increased likelihood of extended use of acetaminophen in pregnancy, but for openness the association was the converse. Given the moderate strength of these associations, it is unlikely that uncontrolled confounding by maternal personality alone would cancel out the association between extended use of acetaminophen in pregnancy and child ADHD found in some studies." Another study, Ystrom et al., 2012[20] came to a similar conclusion and found the association between neuroticism and prenatal use of acetaminophen to be 1.18 (0.91-1.52) with a confidence interval of 95%.

V. Comparing Ji 2018 and Ji 2020

Dr. Pinto-Martin raised concerns about my decision to include Ji et al., 2020[12] and not Ji et al., 2018[21] in my analysis of the literature. Although I *did* include Ji 2018 on my materials considered list, I appreciate the opportunity her comment gives me to reiterate that I had predetermined a priori my approach for including and excluding studies in my Navigation Guide analysis *before* reviewing the published literature. I did that to avoid inadvertently overweighting or underweighting individual studies. Now that the analysis is completed, however, I want to state that including Ji 2018, in addition to Ji 2020 would have, if anything, added *additional* evidence to support my conclusion that prenatal acetaminophen is causally linked to both ADHD and ASD. Therefore, including Ji et al., 2018 would only have bolstered my conclusion, though Ji 2018 remains far less relevant to the question at hand than Ji 2020, because Ji 2018 examined the association between the *mother*'s APAP levels in her bloodstream post-delivery, rather than the levels of APAP in the umbilical cord—a far more relevant measure of prenatal APAP exposure as umbilical cord blood is fetal blood.

I would like to further explain that my approach of selecting the newer report when two or more were available from a single cohort—in the absence of any major strength present in earlier and not in later publications—was driven by the fact that the most recent study would likely have advantages, such as a larger sample size, longer follow-up, or improved methods. Dr. Pinto-Martin's criticism of Ji 2018, which had potential misclassification due to maternal acetaminophen use measured in maternal plasma collected up to 72 hours after delivery, confirms my decision. As between a study measuring a *mother*'s APAP levels and a study measuring *umbilical cord* APAP levels, the umbilical-cord study is quite obviously more relevant to the question at hand.

It is important to clarify that while Dr. Pinto-Martin suggests that the two papers have opposing results, they, in fact, show similar findings for ADHD. Dr. Pinto-Martin fails to mention the study's actual conclusion, which is that "[m]aternal plasma biomarkers of acetaminophen use measured within a few days after delivery were specifically associated with increased risk of ADHD diagnosis in offspring."[21] (Ji 2018). That statement hardly "contradicts" Ji 2020, which concluded that "cord biomarkers of fetal exposure to acetaminophen were associated with significantly increased risk of childhood ADHD." On ADHD in particular, the studies are entirely consistent with one another. This also suggests that the argument regarding maternal postnatal plasma samples reflecting long-term patterns of use is likely valid, even if it leads to a dampened signal.

Nor are the studies actually contradictory on ASD, either. To be sure, the risk ratios are reported in Ji 2020 are *higher* than those in Ji 2018. But that is no surprise given that Ji 2020 employed a superior design for answering the question at issue here, using umbilical cord plasma (rather than the mother's acetaminophen plasma levels post-delivery) as the measurement of exposure. And Ji 2018 still did show positive results even if not statistically significant. As the Ji 2018 authors note, the "ASD diagnosis (without ADHD diagnosis) group[] had more mothers with higher levels of acetaminophen metabolites." As Dr. Pinto-Martin admits in her tables, even Ji 2018 showed the following results *for ASD*: "Total acetaminophen burden: aOR=1.39, 95% CI 0.59-3.27." Although not statistically significant, what that point estimate indicates is that in the Ji 2018 study, women who had higher acetaminophen levels had a 39% increased risk of having a child with ASD. The results from the other models employed to evaluate acetaminophen burden in Ji 2018 show similar results. As the results excerpted from the (actual) Ji 2018 paper make clear, the odds ratio for ASD for women who had detectable levels of acetaminophen burden range from 1.36 on the low end to 1.85 on the high end. Though not statistically significant, these results

are hardly inconsistent with the proposition that there is an association between prenatal APAP exposure and ASD.

Model		ADHD, 188 (15.9%)			ASD, 44 (3.7%)		
		Odds Ratio	95% CI	*p*-Value	Odds Ratio	95% CI	*p*-Value
Acetaminophen burden **							
Model 1	Below median	1.57	(1.05,2.36)	0.029	1.85	(0.87,3.93)	0.110
	Above median	1.88	(1.28,2.78)	0.001	1.94	(0.92,4.08)	0.080
Model 2	Below median	1.56	(1.01,2.42)	0.045	1.39	(0.63,3.06)	0.410
	Above median	1.91	(1.21,3.04)	0.006	1.36	(0.58,3.20)	0.477
Model 3	Below median	1.58	(1.02,2.45)	0.041	1.43	(0.64,3.15)	0.381
	Above median	1.92	(1.21,3.05)	0.006	1.38	(0.59,3.25)	0.457
Model 4	Below median	1.54	(1.00,2.39)	0.052	1.39	(0.63,3.08)	0.411
	Above median	1.89	(1.19,3.01)	0.007	1.37	(0.58,3.22)	0.475
Model 5	Below median	1.56	(1.01,2.41)	0.045	1.39	(0.63,3.07)	0.409
	Above median	1.90	(1.20,3.02)	0.007	1.38	(0.59,3.25)	0.459
Model 6	Below median	1.58	(1.02,2.46)	0.040	1.43	(0.64,3.15)	0.382
	Above median	1.88	(1.18,3.00)	0.008	1.39	(0.59,3.27)	0.456

In sum, though less relevant to the question at issue than Ji 2020, the Ji 2018 study if anything supports my conclusions, rather than refuting them: Women in this study who had higher APAP levels in their bloodstream in general had a higher increased risk of having a child with a NDD.

VI. Errors in Basic Epidemiological Principles

Defendants' experts make a number of erroneous statements regarding epidemiological principles. Their mistakes are actually among the most commonly made, even by well-trained epidemiologists. They have become so common that many epidemiologists, as well as non-epidemiologists, have accepted them as true. This prompted the leading epidemiologists of our time, including Sander Greenland, Stephen Senn, Kenneth Rothman, John Carlin, Charles Poole, Steven Goodman and Douglas Altmanto, to issue a joint statement to stop these continued misstatements of epidemiology.[22] They began by professing that "[m]isinterpretation and abuse of statistical tests, confidence intervals, and statistical power have been decried for decades, yet remain rampant." I address some of those issues below.

a. Advocating for further adjustment for more variables than those included in the available studies is unscientific, unfounded speculation

A minority of the papers that are available in the literature present more than one analysis of the same association between prenatal acetaminophen and the neurodevelopmental outcome of concern. That is, they present more than one analysis each adjusted for a different set of confounders. Dr. Pinto-Martin notes that some of these papers show that the models with the highest number of confounders show smaller—though typically still solidly statistically significant—associations. She imagines then that if there were more confounders to adjust for, the association might disappear. This is pure speculation. Dr. Pinto-Martin doesn't have those data and cannot guess: (1) Whether there is any other measured or unmeasured confounders that has not been collectively considered in the literature; or (2) What the direction of confounding would be and whether the corresponding hypothetical adjustment would direct the association toward or away from the null. As scientists, we must work based on facts, knowledge, and/or methods. Dr. Pinto's reliance on imagination and speculation is not scientifically valid, and not helpful in answering the question in this case.

In addition, based on my experience in working with at least 50+ different epidemiological teams that have resulted in more than 600+ papers, epidemiologists present more than one analysis on the same

association when they believe there is value to show both. Dr. Pinto-Martin seems to imply that the reason for showing multiple sets of analysis is to show that confounders are acting upon the results. In my experience, that is not the case but rather one of the most important reasons is uncertainty as to whether some variables added to the most comprehensive models operate as confounders or intermediates. For instance, birth weight – which Dr. Pinto describes at length in her report – has been suggested to be an intermediate when evaluating the effects of prenatal acetaminophen on child neurodevelopment.[23] If an analysis of prenatal acetaminophen and neurodevelopment is adjusted for birth weight, including birth weight as a covariate would inappropriately bring the risk estimates toward the null, hence creating a misleading result.

It is also worth noting there is immense literature on the pitfalls and biases caused by overadjusting in epidemiology, a commonly accepted tenet that Dr. Pinto-Martin conveniently ignores when making her point.

b. P-Value and Confidence Interval

Dr. Pinto-Martin takes the position that if the confidence interval touches 1.0 or the p-value is >0.05, then the point estimate is "likely due to chance." (Pinto-Martin Report at 12-13). Greenland et al. explain her erroneous position:

> [Misconception] "6. A null-hypothesis P-value greater than 0.05 means that no effect was observed, or that absence of an effect was shown or demonstrated.—No! Observing P > 0.05 for the null hypothesis only means that the null is one among the many hypotheses that have P > 0.05. Thus, unless the point estimate (observed association) equals the null value exactly, it is a mistake to conclude from P > 0.05 that a study found "no association" or "no evidence" of an effect. If the null P-value is less than 1 some association must be present in the data, and one must look at the point estimate to determine the effect size most compatible with the data under the assumed model."

Dr. Pinto-Martin misapplies this principle when discussing Ystrom 2017 and argues that because the confidence intervals for the first and third groups include 1.0, they are not statistically significant and should be ignored. (Pinto-Martin Report at 77). As noted above, statisticians do not apply bright line tests of significance in deciding to consider or fully reject data. A lower confidence interval of 0.99 and 0.96, again as noted above, does not result in the results being disregarded or considered negative. As explained above, this is a situation when it is appropriate to evaluate a one-sided confidence interval. Given the slight insignificance when evaluated as two-sided intervals, these risks would clearly be statistically significant when evaluated as one-sided confidence intervals.

Finally, Dr. Pinto-Martin fails to acknowledge the more detailed analysis on this issue provided by Ystrom et al., 2017:

Table X3. HRs for offspring ADHD By number of days of maternal acetaminophen use during pregnancy.

	All indications	Fever and infections	Pain conditions	Not specified
1-7 days	0.90 (0.81–1.00)	0.90 (0.75–1.09)	0.89 (0.76–1.04)	1.30 (0.98–1.73)
8-14 days	1.18 (0.98–1.42)	1.02 (0.55–1.89)	1.12 (0.83–1.50)	1.96 (1.36–2.82)
15-21 days	1.35 (1.00–1.81)	0.98 (0.24–3.95)	1.43 (0.96–2.14)	1.79 (0.95, 3.35)
22-28 days	1.60 (0.70–3.69)	6.25 (1.71–22.05)	1.08 (0.34–3.39)	-
29 or more	2.20 (1.50–3.24)	2.40 (0.34–16.78)	2.56 (1.54–4.25)	2.13 (0.88–5.15)

This table makes clear that that there is a positive dose-response trend and shows the persistent pattern in her report of selectively considering only the evidence favorable to her otherwise unfunded arguments.

c. Consistency of study results

Dr. Pinto-Martin states that "the results across studies of ASD and/or ADHD are inconsistent," because "[s]ome report statistically significant associations; others do not." (Pinto-Martin Report at 5). This is a glaring methodological error on her part. If some study results show risk ratios above 1.0, and a roughly equal number of study results show risk ratios below 1.0, that pattern may be characterized as inconsistent. But that is not the pattern that Dr. Pinto-Martin is suggesting here for the association between prenatal APAP exposure and NDDs (not could she). Indeed, implicit in her statement above is a concession that the risk ratios in all of these studies are almost uniformly above 1.0—with vanishingly few exceptions. What Dr. Pinto-Martin is instead suggesting is that the results are "inconsistent" due not to the risk ratios, but the confidence intervals. This is a truly basic error, because that is not what "inconsistent" means in this context. Indeed, it is an error cautioned against by basic epidemiology texts. The *Modern Epidemiology* textbook by Rothman et al. is considered a leading authority in our field, and it states (point blank) as follows: "It is a *misconception* to describe studies as 'inconsistent' if some have relative risks with confidence intervals that include 1.0 or p-values greater than 0.05, while others do not." But that is *precisely* what Dr. Pinto-Martin is doing when she suggests that the literature is "inconsistent" simply because it contains a mixture of statistically significant and non-significant results. Her error on this basic point suggests, again, that she is letting her desired results drive the epidemiology rather than the other way around. That error is all the more egregious here given that the vast majority of the results *are* statistically significant. There is simply no question that literature containing dozens of positive results—most of them statistically significant—should be characterized as "consistent" for these purposes. The reason is a fundamental one: a literature like that indicates that in essentially every study, women

who took APAP while pregnant had children with NDDs at higher rates than women who did not take APAP while pregnant. That is a consistent set of results, and one that is compelling evidence in support of causation in this case.

I want to make clear that even if we used the incorrect standards suggested by Dr. Pinto-Martin, the results would still be highly consistent. The overwhelming majority of the studies in the literature shows statistically significant associations between prenatal use of acetaminophen and ADHD and ASD. The argument by Dr. Pinto should not distract from the point that whichever threshold of statistical evidence we considered, the results speak by themselves; the tables I provided in my report's appendix can be used to assess the impressive consistency of the results across studies quickly and unequivocally. I am surprised a fellow epidemiologist would state the opposite.

d. Risk Ratio

Dr. Faraone is wrong that only risk ratios above 3.0 or 4.0 can demonstrate causation. In his report, Dr. Faraone states that risk ratios of "1.21 and 1.34, which are too small to demonstrate causality or rule out confounding." (Faraone Report at 73). He goes on to make the even more sweeping claim that "a risk ratio of less than three or four should not be considered causal." (Faraone Report at 74). He even describes a risk ratio of 3.0 or 4.0 as "the floor for demonstrating a causal association." All of this is simply not true, either theoretically or empirically.

To begin with the theoretical, if a certain exposure in reality undisputedly causes a 20% or 30% increase in a certain risk—as many exposures do—then the observed risk ratios will be 1.2 or 1.3 even in a perfectly designed study. To take a concrete example, randomized clinical trials sometimes show that a certain drug improves outcomes by 20% or 30%. Nobody would argue that these are not causal associations even though the observed risk ratios are 1.2 or 1.3. In sum, there is no theoretical justification for placing arbitrary limits on the magnitude of a risk ratio that can suggest a causal association: sometimes real risks are just less than 3.0 or 4.0. Indeed, Bradford Hill specifically cautioned against *exactly* the kind of reasoning that Dr. Faraone is employing here. As Bradford Hill put it, "[w]e must not be too ready to dismiss a cause-and-effect hypothesis merely on the grounds that the observed association appears to be slight," because "[t]here are many occasions in medicine when this is in truth so."[24] Dr. Faraone falls into exactly the line of erroneous reasoning that Bradford Hill cautions against.

Theory aside, Dr. Faraone's argument flies in the face of reality, because we know for certain that there are many causal relationships that have risk ratios less than 3.0 or 4.0. I detail them extensively in my report, along with supporting citations. They include air pollution and mortality, smoking and heart disease, and secondhand smoker and lung cancer. (Baccarelli Report at 24-26). None of these associations would satisfy Dr. Faraone's arbitrary risk ratio cutoffs, but all of them are considered causal. Put plainly, tobacco companies tried for years to disprove a connection between secondhand smoke and lung cancer using exactly the argument that Dr. Faraone advances. We know now that they were wrong: secondhand smoke causes lung cancer even though the risk ratio is below 3.0 and 4.0.

Finally, it is worth noting that, even if Dr. Faraone was correct about the need for an arbitrary risk ratio cutoff—and he is not correct—there is every reason to think that the true risk ratio is higher than 1.2 or 1.3. The studies that produced those risk estimates relied on self-reporting of acetaminophen exposure, a rough proxy that almost certainly led to non-differential misclassification and bias toward the null. In other words, those studies likely *understated* the true risk. This is evident from the fact that studies with better exposure measures—like Ji and Baker, which specifically measured the amount of acetaminophen in fetal umbilical plasma and meconium—produced much higher risk ratios. Indeed, those studies

produced risk ratios that would satisfy even Dr. Faraone's (entirely erroneous) requirement for a risk ratio cutoff before considering an association causal.

VII. The Navigation Guide is a valid methodology for assessing the causal relationship between prenatal use of acetaminophen and ASD and ADHD

Dr. Pinto-Martin is wrong when she states the Navigation Guide is not a "valid methodology" because it is "not designed to be used . . . to evaluate a causation connection between an exposure and outcome." (Pinto-Martin Report at 110-111). To the contrary, determining whether there is a causal relationship between exposure and adverse outcome is fundamental to the Navigation Guide inquiry and underpins Steps 3 and 4 (grading strength of each study and rendering opinion on overall quality of evidence) of the methodology.[25]

Dr. Pinto-Martin provides no authority for her conclusion that the Navigation Guide "is a poor fit for assessing whether prenatal acetaminophen use causes ASD and ADHD." (Pinto-Martin Report at 111-112). The Navigation Guide is based on the GRADE system (Grading of Recommendations Assessment, Development and Evaluation), which "systematically rates the quality of the evidence" and is widespread in clinical review to assess causal associations.[25] ("GRADE is also in wide use, having been adopted by over 50 organizations, including the World Health Organization, AHRQ, CDC and Kaiser Permanente"). The Navigation Guide differs from GRADE in that it was created to address the differences in the types of evidence available to assess environmental hazards in that environmental contaminants and other adverse exposures do not have required regulatory testing and do not have randomized control trials (RCTs), as is usually the case with clinical science.[25] That is the same situation we are faced with here as acetaminophen did not have required regulatory testing before it was brought to market— any pre-clinical testing would have been too short in duration to detect this risk in any event—and RCTs regarding prenatal use of acetaminophen are no longer possible due to the ethical concerns associated with such testing. For both environmental contaminants and acetaminophen, the evidence is limited to observational and preclinical studies, and the Navigation Guide is therefore uniquely suited to assess the causal relationship between acetaminophen and the neurodevelopmental disorders of ASD and ADHD.

Contrary to what Dr. Pinto-Martin suggested, the Navigation Guide does not require application of the precautionary principle or to apply a lower causation threshold. (Pinto-Martin Report at 111). It is simply a framework to assess scientific evidence that does not over-emphasize randomized control trials, which are often not available tools to assess causal associations. I did not in any way apply a lower standard of causal association in my application of the Navigation Guide to the evidence I reviewed.

Nonetheless, Dr. Pinto-Martin's criticism as to whether the Navigation Guide can assess causation does not ultimately matter because the Navigation Guide simply provides for an objective framework to systematically assess and <u>transparently</u> document scientific evidence. As the name states, the Navigation Guide is a "guide" and provides transparency in that it requires scientists to follow and document certain steps and objectively analyze each study and then transparently rate each study considered as part of the causal analysis.[26] While following predetermined steps, the Navigation Guide requires using tables (included in Appendix 1 of my report) to document the entirety of the literature reviews, the strengths and weaknesses of each study considered in the analysis, and the strength of evidence derived from each study. By using the Navigation Guide, I was able to objectively review each piece of epidemiological evidence and document and explain the strength I assigned each study, which I also incorporated into my Bradford Hill analysis. As a result, anyone can see the weight I assigned to each study for my Bradford Hill analysis. Defendants' experts, including Dr. Pinto-Martin, Dr. Faraone, and Dr. Kolevzon, all notably fail to

objectively rate the studies when assessing the epidemiological evidence. As a result, while I made my evaluation of every single study that I reviewed available, documented, and transparent, their assessments of each study seem subjective and opaque.

VIII. My reliance on studies without ASD and ADHD clinical diagnoses as endpoints is proper

Defendants' experts criticize that I rely on studies that do not have clinical diagnoses of ASD and ADHD and that I rely on studies with endpoints of general neurodevelopmental disorder symptoms. I rely on and incorporate my colleague, Dr. Eric Hollander's rebuttal report, as Dr. Hollander effectively and thoroughly explains how these criticisms are unfounded. To be sure, studies showing an association between prenatal APAP exposure and clinical *diagnoses* of ADHD and ASD provide *particularly* powerful evidence of a causal relationship. But the power of those studies in no way suggests that researchers should blind themselves to the results from studies using other endpoints that are obviously relevant to the question at hand. For example, Dr. Pinto-Martin criticizes me for relying on a study (Avella-Garcia 2016[27]) that used results from the Childhood Autism Spectrum Test (CAST) as an endpoint. (Pinto Martin Report at 48). But surely a researcher evaluating whether prenatal APAP exposure causes autism would want to at least consider the fact that a study showed that women who took APAP while pregnant were more likely to have children displaying signs of autism on a test (CAST) specifically designed to look for signs of autism. Dr. Pinto-Martin never explains why it would be unreasonable to consider that kind of evidence—or indeed why it would be reasonable not to do so.

IX. Conclusion

In conclusion, I disagree and rebut the points outlined here that have been asserted by Defendants' experts, particularly Dr. Pinto-Martin. As set forth in my initial report, and further supported by the points I make here, there is a causal relationship between prenatal use of acetaminophen and the neurodevelopmental disorders of ASD and ADHD. My opinion is made to a reasonable degree of medical, epidemiological, and scientific certainty.

1. Ystrom E, Gustavson K, Brandlistuen RE, et al. Prenatal Exposure to Acetaminophen and Risk of ADHD. *Pediatrics*. 2017;140(5):9. doi:10.1542/peds.2016-3840

2. Liew Z, Kioumourtzoglou MA, Roberts AL, O'Reilly ÉJ, Ascherio A, Weisskopf MG. Use of Negative Control Exposure Analysis to Evaluate Confounding: An Example of Acetaminophen Exposure and Attention-Deficit/Hyperactivity Disorder in Nurses' Health Study II. *Am J Epidemiol*. 2019;188(4):768-775. doi:10.1093/aje/kwy288

3. Stergiakouli E, Thapar A, Davey Smith G. Association of Acetaminophen Use During Pregnancy With Behavioral Problems in Childhood: Evidence Against Confounding. *JAMA Pediatr*. 2016;170(10):964. doi:10.1001/jamapediatrics.2016.1775

4. Trønnes JN, Wood M, Lupattelli A, Ystrom E, Nordeng H. Prenatal paracetamol exposure and neurodevelopmental outcomes in preschool-aged children. *Paediatr Perinat Epidemiol*. 2020;34(3):247-256. doi:10.1111/ppe.12568

5. Rossitto M, Marchive C, Pruvost A, et al. Intergenerational effects on mouse sperm quality after in utero exposure to acetaminophen and ibuprofen. *FASEB J Off Publ Fed Am Soc Exp Biol*. 2019;33(1):339-357. doi:10.1096/fj.201800488RRR

6. Eslamimehr S, Jones AD, Anthony TM, et al. Association of prenatal acetaminophen use and acetaminophen metabolites with DNA methylation of newborns: analysis of two consecutive generations of the Isle of Wight birth cohort. *Environ Epigenetics*. 2022;8(1):dvac002. doi:10.1093/eep/dvac002

7. Ricci C, Albanese CM, Pablo LA, et al. In utero acetaminophen exposure and child neurodevelopmental outcomes: Systematic review and meta-analysis. *Paediatr Perinat Epidemiol*. 2023;n/a(n/a). doi:10.1111/ppe.12963

8. Gustavson K, Ystrom E, Ask H, et al. Acetaminophen use during pregnancy and offspring attention deficit hyperactivity disorder – a longitudinal sibling control study. *JCPP Adv*. 2021;1(2):e12020. doi:10.1002/jcv2.12020

9. Sjölander A, Zetterqvist J. Confounders, Mediators, or Colliders: What Types of Shared Covariates Does a Sibling Comparison Design Control For? *Epidemiol Camb Mass*. 2017;28(4):540-547. doi:10.1097/EDE.0000000000000649

10. Brandlistuen RE, Ystrom E, Nulman I, Koren G, Nordeng H. Prenatal paracetamol exposure and child neurodevelopment: a sibling-controlled cohort study. *Int J Epidemiol*. 2013;42(6):1702-1713. doi:10.1093/ije/dyt183

11. Rothman K, Greenland S, Lash T. *Modern Epidemiology*. 3rd. Edition. Lippincott Williams & Wilkins

12. Ji Y, Azuine RE, Zhang Y, et al. Association of Cord Plasma Biomarkers of In Utero Acetaminophen Exposure With Risk of Attention-Deficit/Hyperactivity Disorder and Autism Spectrum Disorder in Childhood. *JAMA Psychiatry*. 2020;77(2):180. doi:10.1001/jamapsychiatry.2019.3259

13. Bai D, Yip BHK, Windham GC, et al. Association of Genetic and Environmental Factors With Autism in a 5-Country Cohort. *JAMA Psychiatry*. 2019;76(10):1035-1043. doi:10.1001/jamapsychiatry.2019.1411

14. Modabbernia A, Velthorst E, Reichenberg A. Environmental risk factors for autism: an evidence-based review of systematic reviews and meta-analyses. *Mol Autism*. 2017;8(1):13. doi:10.1186/s13229-017-0121-4

15. Boomsma DI, Saviouk V, Hottenga JJ, et al. Genetic Epidemiology of Attention Deficit Hyperactivity Disorder (ADHD Index) in Adults. *PLOS ONE*. 2010;5(5):e10621. doi:10.1371/journal.pone.0010621

16. Larsson H, Chang Z, D'Onofrio BM, Lichtenstein P. The heritability of clinically diagnosed Attention-Deficit/Hyperactivity Disorder across the life span. *Psychol Med*. 2014;44(10):2223-2229. doi:10.1017/S0033291713002493

17. Moore DS, Shenk D. The heritability fallacy. *Wiley Interdiscip Rev Cogn Sci*. 2017;8(1-2). doi:10.1002/wcs.1400

18. Qiu S, Qiu Y, Li Y, Cong X. Genetics of autism spectrum disorder: an umbrella review of systematic reviews and meta-analyses. *Transl Psychiatry*. 2022;12:249. doi:10.1038/s41398-022-02009-6

19. Lupattelli A, Trinh NTH, Nordeng H. Association of maternal personality traits with medication use during pregnancy to appraise unmeasured confounding in long-term pharmacoepidemiological

safety studies. *Front Pharmacol*. 2023;14. Accessed July 27, 2023. https://www.frontiersin.org/articles/10.3389/fphar.2023.1160168

20. Ystrom E, Vollrath ME, Nordeng H. Effects of personality on use of medications, alcohol, and cigarettes during pregnancy. *Eur J Clin Pharmacol*. 2012;68(5):845-851. doi:10.1007/s00228-011-1197-y

21. Ji Y, Riley AW, Lee LC, et al. Maternal Biomarkers of Acetaminophen Use and Offspring Attention Deficit Hyperactivity Disorder. *Brain Sci*. 2018;8(127):15. doi:doi:10.3390/brainsci8070127

22. Greenland S, Senn SJ, Rothman KJ, et al. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *Eur J Epidemiol*. 2016;31:337-350. doi:10.1007/s10654-016-0149-3

23. Arneja J, Hung RJ, Seeto RA, et al. Association between maternal acetaminophen use and adverse birth outcomes in a pregnancy and birth cohort. *Pediatr Res*. 2020;87(7):1263-1269. doi:10.1038/s41390-019-0726-8

24. Hill AB. The Environment and Disease: Association or Causation? *Proc R Soc Med*. 1965;58(5):295-300.

25. Woodruff TJ, Sutton P, The Navigation Guide Work Group. An Evidence-Based Medicine Methodology To Bridge The Gap Between Clinical And Environmental Health Sciences. *Health Aff (Millwood)*. 2011;30(5):931-937. doi:10.1377/hlthaff.2010.1219

26. Woodruff TJ, Sutton P. The Navigation Guide Systematic Review Methodology: A Rigorous and Transparent Method for Translating Environmental Health Science into Better Health Outcomes. *Environ Health Perspect*. 2014;122(10):1007-1014. doi:10.1289/ehp.1307175

27. Avella-Garcia CB, Julvez J, Fortuny J, et al. Acetaminophen use in pregnancy and neurodevelopment: attention function and autism spectrum symptoms. *Int J Epidemiol*. 2016;45(6):dyw115. doi:10.1093/ije/dyw115

I hold the foregoing opinions to a reasonable degree of scientific and medical certainty.

Dated: July 28, 2023

Respectfully submitted,

Type text here

Andrea Baccarelli, MD, PhD, MPH